VALUE IN HEALTH 15 (2012) 495–503

Available online at www.sciencedirect.com
SciVerse ScienceDirect
journal homepage: www.elsevier.com/locate/jval

Mapping the Functional Assessment of Cancer Therapy-General or -Colorectal to SF-6D in Chinese Patients with Colorectal Neoplasm

Carlos K.H. Wong, MPhil, BSc[1,], Cindy L.K. Lam, MBBS, MD, FRCGP, FHKAM (Family Medicine)[1], Donna Rowen, PhD[2], Sarah M. McGhee, PhD, BSc, BA, Hon. MFPH (RCP UK)[3], Ka-Ping Ma, BSc[1], Wai-Lun Law, MBBS, MS, FRCS(Ed), FCSHK, FHKAM (Surgery), FACS[4], Jensen T.C. Poon, MBBS, MS, FRCSEd(Gen), FHKAM (Surgery)[4], Pierre Chan, MBBS, FHKAM (Medicine)[5], Dora L.W. Kwong, MBBS, MD, FRCR, FHKCR, FHKAM[6], Janice Tsang, MBBS, MRCP(UK), FRCP(Edin), FHKCP, FHKAM (Medicine)[6]*

[1]*Department of Family Medicine and Primary Care, The University of Hong Kong, Hong Kong;* [2]*Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, UK;* [3]*Department of Community Medicine, School of Public Health, The University of Hong Kong, Hong Kong;* [4]*Department of Surgery, The University of Hong Kong, Hong Kong;* [5]*Department of Medicine, Queen Mary Hospital, Hong Kong;* [6]*Department of Clinical Oncology, The University of Hong Kong, Hong Kong*

ABSTRACT

Objectives: To map Functional Assessment of Cancer Therapy-General (FACT-G) and Functional Assessment of Cancer Therapy-Colorectal (FACT-C) subscale scores onto six-dimensional health state short form (derived from short form 36 health survey) (SF-6D) preference-based values in patients with colorectal neoplasm, with and without adjustment for clinical and demographic characteristics. These results can then be applied to studies that have used FACT-G or FACT-C to predict SF-6D utility values to inform economic evaluation.
Methods: Ordinary least square regressions were estimated mapping FACT-G and FACT-C onto SF-6D by using cross-sectional data of 537 Chinese subjects with different stages of colorectal neoplasm. Mapping functions for SF-6D preference-based values were developed separately for FACT-G and FACT-C in four sequential models for addition of variables: 1) main-effect terms, 2) squared terms, 3) interaction terms, and 4) clinical and demographic variables. Predictive performance in each model was assessed by the R^2, adjusted R^2, predicted R^2, information criteria (Akaike information criteria and Bayesian information criteria), the root mean square error, the mean absolute error, and the proportions of absolute error within the threshold of 0.05 and 0.10.
Results: Models including FACT variables and clinical and demographic variables had the best predictive performance measured by using R^2 (FACT-G: 59.98%; FACT-C: 60.43%), root mean square error (FACT-G: 0.086; FACT-C: 0.084), and mean absolute error (FACT-G: 0.065; FACT-C: 0.065). The FACT-C–based mapping function had better predictive ability than did the FACT-G–based mapping function.
Conclusions: Models mapping FACT-G and FACT-C onto SF-6D reached an acceptable degree of precision. Mapping from the condition-specific measure (FACT-C) had better performance than did mapping from the general cancer measure (FACT-G). These mapping functions can be applied to FACT-G or FACT-C data sets to estimate SF-6D utility values for economic evaluation of medical interventions for patients with colorectal neoplasm. Further research assessing model performance in independent data sets and non-Chinese populations are encouraged.

Keywords: colorectal neoplasm, FACT-C, mapping, patient-reported outcomes, quality of life, SF-6D.

Introduction

Colorectal neoplasm (CRN) is becoming a leading burden of disease worldwide [1]. More advanced clinical interventions including screening and chemotherapeutic treatment have been developed for CRN, but most of them are costly. Economic evaluation using cost-effectiveness analysis is commonly used to appraise clinical interventions to inform resource allocation. Quality-adjusted life-years are used as a general measure of health effects to capture both morbidity and mortality [2]. The "Q" quality adjustment weight of the quality-adjusted life-year, ranging from 0 for death to 1 for perfect health, is calculated by using a preference-based valuation of health elicited by direct or indirect methods [3]. Direct valuation techniques, such as standard gamble or time trade-off of colorectal cancer (CRC) health states [4], can be used to elicit utility values from patients for their own health, but its validity is questioned [5,6] on the grounds that utility should be judged by society as a whole. An alternative method is the use of preference-based measures that have been valued by the general population, for example, the EuroQol five-dimensional (EQ-5D) questionnaire [7] and the six-dimensional health state short form (derived from short form 36 health survey (SF-6D) [8,9]. The EQ-5D questionnaire is recommended by the National Institute of Health and Clinical Excellence [2] and is widely used in Europe. The SF-6D

* *Address correspondence to:* Carlos K.H. Wong, Department of Family Medicine and Primary Care, The University of Hong Kong, 3/F, Ap Lei Chau Clinic, 161 Ap Lei Chau Main Street, Ap Lei Chau, Hong Kong.
E-mail: carlosho@hku.hk.

1098-3015/$36.00 – see front matter Copyright © 2012, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.
doi:10.1016/j.jval.2011.12.009

derived from health states measured by Medical Outcomes Study SF-36 Health Survey [8] is more popular in the United States and Asia. Population-specific scoring algorithms for SF-6D are available for the United Kingdom [8,9], Japan [10], Portugal [11], Hong Kong [12], and Brazil [13].

Condition-specific health-related quality-of-life (HRQOL) measures are usually not preference based. "Mapping" the generic or condition-specific HRQOL measures onto generic preference-based measures is a method that can be used to predict utility values even though the preference-based measure was not included in the study. Mapping involves two stages. First, regression analysis is used to estimate the relationship between the "source" non–preference-based measure and the "target" preference-based measure. Second, this relationship can be applied to a study data set containing the source measure to predict. In recent years, numerous studies have successfully developed functions mapping from non–preference-based instruments onto EQ-5D questionnaire and SF-6D preference-based values in different patient groups [14–16]; see Brazier et al. [17] for a recent overview. Most studies in the literature have mapped onto the EQ-5D questionnaire [17]. Yet studies mapping onto both the EQ-5D questionnaire and SF-6D found that the models mapping onto SF-6D had better predictive performance than did those mapping onto the EQ-5D questionnaire [14,15].

The Functional Assessment of Cancer Therapy-Colorectal (FACT-C) is a popular HRQOL measure specific for CRN. Its psychometric properties were tested and confirmed in Chinese patients with colorectal polyps or cancer [18]. The FACT-C is an extended version of the FACT-General (FACT-G), which makes up four of the five subscales of the former. Mappings from FACT-G or FACT-C onto SF-6D provide a useful function for converting profile HRQOL scores collected by these condition-specific measures commonly used in cancer clinical trials to preference-based values for the calculation of quality-adjusted life-years for cost-effectiveness analysis.

Mapping studies often incorporate demographic and clinical characteristics into model estimation to increase a model's predictive performance [16,19–21]. Demographics (age and gender) and tumor stage of CRN variables were included in the analysis. The aim of this study was to map from FACT-G and FACT-C subscale scores to SF-6D preference-based values in patients with CRN, with and without adjustment for clinical and demographic characteristics. The mapping function allowed researchers to conduct health economic appraisals of preventive screening and treatment program for CRN.

Methods

Subjects

All Chinese patients 18 years or older with a histology of confirmed colorectal polyp or cancer for at least 6 months were recruited from specialist outpatient medical and surgical colorectal clinics of a regional hospital in Hong Kong between October 2009 and September 2010. Patients were excluded if the doctor judged their life expectancy to be less than 6 months; they were unable to communicate in Cantonese; they had a known cognitive impairment; or they were too ill to participate in an interview. A total of 575 patients were invited to participate and of these 20 were unreachable and 2 withdrew, meaning 553 subjects were successfully interviewed via telephone (449 subjects) or face to face (104 subjects) by trained interviewers. Clinical characteristics such as tumor stage, treatment status, and colostomy status were retrieved from the medical records of patients. The CRN was classified into six stages on the basis of screening surveillance guideline [22] and American Joint Committee on Cancer (AJCC) staging system [23]: 1) low-risk polyps group (patients with ≤2 adenomas or 3–4 adenomas all of which were not larger than 1 cm), 2) high-risk polyps group (patients with ≥5 adenomas or with ≥3 adenomas at least one of which was larger than 1 cm), 3) stage I CRC, 4) stage II CRC, 5) stage III CRC, and 6) stage IV CRC. The tumor stage of CRN was classified as unknown if it was not specified in the medical record. Demographic and clinical characteristics of the patients are presented in Table 1. Patients had an average age of 63.2 years (SD: 11.3) and 42.0% were females. For the stage of CRN, the mode was "stage III" with a percentage of 20.6%; the proportions of the other stages ranged from 13.0% to 18.3%. Among all patients, eight subjects (1.4%) had unknown stage of diagnosis and were excluded before the regression analysis of the mapping functions of SF-6D.

Table 1 – Sociodemographic and clinical characteristics of study subjects.

	Total (n = 553)
Age (y), mean ± SD	63.2 ± 11.3
Sex, n (%)	
Male	321 (58.0)
Female	232 (42.0)
Stage of colorectal neoplasm, n (%)	
Low-risk polyp	93 (16.8)
High-risk polyp	72 (13.0)
Stage I	83 (15.0)
Stage II	101 (18.3)
Stage III	114 (20.6)
Stage IV	82 (14.8)
Unknown	8 (1.4)
Duration of diagnosis (mo), mean ± SD	46.6 ± 55.8
Treatment status, n (%)*	
Palliative	63 (16.4)
Adjuvant	26 (6.8)
No	296 (76.9)
Stoma (%)*	
Yes	51 (13.2)
No	334 (86.8)

* Colorectal cancer patients only (n = 385).

Data collection

Preference-based value and condition-specific HRQOL data were collected by using the Hong Kong Chinese version of SF-6D and traditional Chinese (Hong Kong) FACT-C Version 4, respectively. These instruments were administrated by trained interviewers at the same time. Clinical data were obtained from the doctors or extracted from the medical records of patients [24]. This study was approved by institutional review boards at the University of Hong Kong and Hospital Authority Hong Kong West Cluster (IRB Ref# UW 09-391).

Study instruments

FACT-G, developed by the Center on Outcomes, Research and Education [25], is widely used to measure HRQOL for cancer patients in clinical trials. It is the core member of FACT instruments including four subscales: 1) 7-item physical well-being (PWB), 2) 7-item social/family well-being (SWB), 3) 6-item emotional well-being (EWB), and 4) 7-item functional well-being (FWB). FACT-C (Version 4) is the CRC-specific member of FACT instruments that includes the four FACT-G subscales plus a 9-item colorectal cancer subscale (CCS) that addresses additional concerns on quality of life for CRC [26]. All items have five response options ("not at all," "a little bit," "somewhat," "quite a bit," and "very much"). Each domain has a subscale score where higher scores indicate better HRQOL. Both FACT-G and FACT-C are designed for either patient self-completion [25] or interviewer administration [18,27]. They can be admin-

istered face to face or on telephone by trained interviewers [25]. The traditional Chinese version of FACT-C has been shown to be valid and reliable in Chinese patients with colorectal polyps or CRC in Hong Kong [18].

The SF-6D is a preference-based measure of health consisting of six aspects of health dimensions—physical functioning, role limitations, social functioning, pain, mental health, and vitality; each dimension is classified into four to six levels, giving a total of 18,000 combinations of health states. The SF-6D was originally developed by Brazier et al.[8]. It was translated into the Hong Kong Chinese version by Lam et al. [12], which was validated and valued in the Hong Kong general Chinese population by using standard gamble [12,28]. The Hong Kong population-specific preference scoring algorithm was used in this study here because these preferences are most representative of this patient population. The Hong Kong Chinese version of the SF-6D can be administrated by self or an interviewer [29]. Telephonic interview was acceptable to administer SF-6D because our preliminary analysis showed that the SF-6D values difference between telephone (449 subjects) and face-to-face (104 subjects) interviews was not statistically significant ($P = 0.152$) by nonparametric Mann-Whitney U test.

Statistical analysis

Model specification

Model specification was informed by prior descriptive analysis of all variables. Separate models were estimated by mapping FACT-C and FACT-G onto SF-6D. Models were estimated by using 1) FACT variables using main effects, 2) as per point 1 plus squared terms, 3) as per point 1 plus interaction terms, and 4) as per point 3 plus clinical and demographics variables. Additional variables were added by using the stepwise regression technique. Ordinary least square method was used for all analysis. The F test was used to retain variables with an inclusion criterion of P value lower than 0.05 and exclusion criterion of P value greater than 0.10. Lower-order main effect terms (first order) were retained in the model even if their effects became insignificant when squared and interaction terms (second order) were added because of the principle of hierarchy. To ensure comparability of different models, only subjects with complete data ($n = 537$) across all variables were used in the analysis.

Model comparison

R^2 and the adjusted R^2 (Adj R^2) statistics were used to measure the explanatory power of each model, that is, how much of the variability in the dependent variable was explained by the predictors. In addition, predicted R^2 (Pred R^2) was calculated by one minus a fraction of predicted residual sum of squares over total sum of squares, which is more able than R^2 to assess the performance of a model for predicting new observations [30]. Of prime importance for selecting between models is predictive performance. Models with low errors are preferred because this implies that the model will have better predictive performance when used in other data sets to predict SF-6D. To assess the predictive ability of a model, the differences between the predicted and observed SF-6D values at the individual level were examined by computing the root mean square error (RMSE) and the mean absolute error (MAE). In addition, the number of observations and the corresponding proportions in the full sample that the absolute error (AE) was greater than the threshold of 0.05 and 0.10 for each individual were calculated, respectively. The goodness of fit of the model was described by using Akaike information criteria and Bayesian information criteria statistic. The achievable ranges of the SF-6D preference-based values from the resulting mapping functions were compared with the observed range of the SF-6D values. Residual plots for the final models were used as the primary tool to examine model adequacy such as the non-normality and heteroscedasticity.

Model validation

The validity and robustness of the mapping functions were assessed by using the 10-fold cross-validation procedure [20]. The full sample was randomly split into 10 equally sized groups. Each combination of nine groups formed a training data set that was used to estimate the parameters of the regression model, while the remaining group was considered as a test data set that examined the errors of prediction of the model generated by the training data set. The test data set was validated by inputting the independent variables of each observation of the test data set into the model fitted by the training data set. Differences of the observed value of the dependent variable from the predicted value calculated from the model were the error of prediction. The above procedure was repeated until all the 10 possible training data sets were tested. The errors of prediction for all observations were obtained to calculate the RMSE, MAE, and percentages of AEs greater than 0.05 and 0.10 ($AE > 0.05$ and $AE > 0.10$).

All regressions and other statistical analyses were conducted by using the SPSS, Version 18.0 (SPSS, IBM, Inc., Chicago, IL).

Results

Descriptive statistics for SF-6D and FACT-C

Observed SF-6D and FACT-C/FACT-G subscale scores of the subjects are shown in Table 2. The mean score of SF-6D was 0.825 (SD:

Table 2 – Descriptive statistics for SF-6D preference-based values and FACT-C subscale scores.

	Mean ± SD	95% CI	%Floor	%Ceiling	Observed range	Theoretical range
Preference-based values						
SF-6D	0.825 ± 0.136	0.813–0.836	0.00	4.28	0.385–1.00	0.315–1.00
FACT-C subscale						
PWB	25.69 ± 3.25	25.41–25.96	0.00	40.22	4–28	0–28
SWB	19.89 ± 4.36	19.52–20.26	0.37	3.72	0–28	0–28
EWB	21.34 ± 2.90	21.09–21.58	0.00	9.87	9–24	0–24
FWB	18.92 ± 4.38	18.55–19.29	0.00	2.61	1–28	0–28
CCS	21.74 ± 3.14	21.47–22.00	0.00	2.05	9–28	0–28

Note. PWB, SWB, EWB, and FWB form the FACT-G.
CCS, colorectal cancer subscale; CI, confidence interval; EWB, emotional well-being; FACT-C, Functional Assessment of Cancer Therapy-Colorectal; FACT-G, Functional Assessment of Cancer Therapy-General; FWB, functional well-being; PWB, physical well-being; SF-6D, six-dimensional health state short form (derived from short form 36 health survey); SWB, social/family well-being.

Table 3 – Prediction models for patients with colorectal neoplasm using main effects, squared terms of FACT-G, and clinical and demographical variables.

	FACT-G–based model					
	Main effects (G1)		Squared terms added (G2)		Clinical and demographical variables added (G3)	
	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI
PWB	0.01770*	(0.01469,0.02071)	−0.02526*	(−0.04265,−0.00787)	−0.02705*	(−0.04434,−0.00976)
EWB	0.00446*	(0.00119,0.00773)	−0.02439	(−0.05149,0.00270)	−0.02356	(−0.05052,0.00339)
FWB	0.01128*	(0.00917,0.01338)	0.00274	(−0.00485,0.01033)	0.00324	(−0.00431,0.01078)
PWB^2			0.00100*	(0.00061,0.00139)	0.00102*	(0.00063,0.00140)
EWB^2			0.00076*	(0.00004,0.00147)	0.00072*	(0.00001,0.00143)
FWB^2			0.00024*	(0.00003,0.00045)	0.00023*	(0.00001,0.00044)
Stage†						
Low risk					0.04238*	(0.01528,0.06947)
High risk					0.01294	(−0.01574,0.04163)
Stage I					0.02808*	(0.00031,0.05586)
Stage II					0.03908*	(0.01254,0.06562)
Stage III					0.02692*	(0.00133,0.05251)
Female					−0.01963*	(−0.03523,−0.00403)
Constant	0.06169	(−0.00643,0.12981)	0.82992*	(0.55594,1.10390)	0.84470*	(0.57168,1.11771)
R^2	54.95%		58.56%		59.98%	
Adj R^2	54.70%		58.10%		59.07%	
Pred R^2	53.74%		56.70%		57.22%	
AIC	−1041.27		−1080.18		−1086.87	
BIC	−1019.84		−1045.89		−1026.87	
RMSE	0.091		0.087		0.086	
MAE	0.070		0.067		0.067	
AE > 0.05	300 (55.87%)		296 (55.12%)		277 (51.58%)	
AE > 0.10	122 (22.72%)		122 (22.72%)		128 (23.84%)	
Range of fitted	0.299–0.980		0.490–1.023		0.491–1.038	

Adj R^2, adjusted R^2; AE, absolute error; AIC, Akaike information criteria; BIC, Bayesian information criteria; CI, confidence interval; fitted, fitted values; EWB, emotional well-being; FACT-G, Functional Assessment of Cancer Therapy-General; FWB, functional well-being; MAE, mean absolute error; Pred R^2, predicted R^2; PWB, physical well-being; RMSE, root mean square error.

* Significant with $P < 0.05$.

† Stage IV is the reference category of colorectal neoplasm staging.

0.136, 95% confidence interval: 0.813–0.836), ranging from 0.385 to 1, which was narrower than the theoretical range of 0.315 to 1 based on Hong Kong data [28]. There was no floor effect in measuring the SF-6D score in the sample while a slight ceiling effect was observed, with 4.28% of patients having the best possible level of health state value of 1. There were no floor effects observed in the subscales except a very small one in SWB (0.37%). Ceiling effects were observed in all FACT-C subscales, with the highest found in PWB (40.22%).

Regression modeling of the FACT-G–based mapping function

Table 3 shows the results of the analyses with FACT-G subscales scores. The main-effect terms selected from the FACT-G subscale scores in the first model (model G1) were PWB, EWB, and FWB. All the squared terms of PWB, EWB, and FWB scores were included in the second model (model G2) whereas their no interaction terms were significant (model not shown). Tumor stage and sex were eventually added to the model in the last model (model G3). Although PWB, EWB, and FWB contributed significantly in the first model (model G1), the effects of the main-effect terms of EWB and FWB became insignificant after the squared terms were added in the second model (model G2). Sex but not age was significant in the final model (model G3). The effects of tumor stage were significant except the subgroup effect of "high-risk polyps." The final model had a R^2 of 59.98%, an Adj R^2 of 59.07%, and a Pred R^2 of 57.22%. RMSE and MAE were 0.086 and 0.067, respectively. The percentage of AE greater than 0.05 was 51.58% while that of AE greater than 0.10 was 23.84%. The fitted range of the SF-6D preference-based values estimated for the mapping function was 0.491 to 1.038, the lower bound of which was 0.11 more than that of the observed SF-6D range (0.385–1). The scatter plot of mean observed and predicted SF-6D preference-based values by the health state ranking in descending order of the mean observed values (Table 5 and Fig. 1) indicates overprediction for more severe health states, where observed SF-6D value was less than 0.8.

The performance of model G2 without the demographic and clinical variables was only slightly inferior to that of the final model, with an adjusted R^2 of 58.1%, and favorable RMSE, MAE, and AE proportions.

Regression modeling of the FACT-C–based mapping function

Mapping function of FACT-C subscales scores is shown in Table 4. Similar to the model for FACT-G, PWB, EWB, and FWB scores were selected as the main-effect terms in first model (model C1) to which was added the CCS. Among the four squared terms, PWB2 and EWB2 entered the model in the second model (model C2). EWB2, however, was removed from the model in the third model (model C3) that integrated the interaction terms of PWB × FWB, EWB × CCS, and FWB × CCS. In the final model (model C4), tumor effects of "high-risk polyps," "stage I CRC," and "stage III CRC" were insignificant. The final model achieved a R^2 of 61.65%, an Adj R^2 of 60.63%, a Pred R^2 of 58.41%, and the smallest Akaike information criteria among prediction models. RMSE and MAE

Table 4 – Prediction models for patients with colorectal neoplasm using main effects, squared terms, interaction terms of FACT-C, and clinical and demographical variables.

	FACT-C–based model							
	Main effects (C1)		Squared terms added (C2)		Interaction terms added (C3)		Clinical and demographical variables added (C4)	
	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI
PWB	0.01651*	(0.01343,0.01959)	−0.02731*	(−0.04456,−0.01007)	−0.02091*	(−0.03820,−0.00362)	−0.02318*	(−0.04040,−0.00595)
EWB	0.00352*	(0.00022,0.00682)	−0.02741*	(−0.05411,−0.00071)	−0.03780*	(−0.05613,−0.01947)	−0.03530*	(−0.05361,−0.01699)
FWB	0.01028*	(0.00810,0.01338)	0.01025*	(0.00814,0.01237)	0.01442*	(0.00080,0.02804)	0.01455*	(0.00105,0.02805)
CCS	0.00479*	(0.00174,0.00784)	0.00379*	(0.00081,0.00676)	−0.01911*	(−0.03516,−0.00306)	−0.01709*	(−0.03307,−0.00111)
PWB^2			0.00101*	(0.00063,0.00140)	0.00059*	(0.00016,0.00103)	0.00065*	(0.00022,0.00109)
EWB^2			0.00082*	(0.00011,0.00152)				
PWB × FWB					0.00076*	(0.00025,0.00126)	0.00067*	(0.00016,0.00117)
EWB × CCS					0.00200*	(0.00112,0.00289)	0.00185*	(0.00096,0.00273)
FWB × CCS					−0.00104*	(−0.00158,−0.00049)	−0.00095*	(−0.00149,−0.00040)
Stage†								
Low risk							0.03731*	(0.01040,0.06421)
High risk							0.00896	(−0.01937,0.03729)
Stage I							0.02641	(−0.00091,0.05373)
Stage II							0.03581*	(0.00963,0.06198)
Stage III							0.02348	(−0.00193,0.04889)
Female							−0.01921*	(−0.03451,−0.00391)
Constant	0.02703*	(−0.04405,0.09812)	0.77787*	(0.49933,1.05641)	1.03518*	(0.72557,1.34479)	1.01899*	(0.71030,1.32767)
R^2	55.74%		58.67%		60.43%		61.65%	
Adj R^2	55.41%		58.20%		59.83%		60.63%	
Pred R^2	54.31%		56.81%		58.07%		58.41%	
AIC	−1048.81		−1081.50		−1100.90		−1105.78	
BIC	−1023.09		−1047.22		−1058.04		−1037.21	
RMSE	0.090		0.087		0.085		0.084	
MAE	0.069		0.067		0.065		0.065	
AE > 0.05	302 (56.24%)		273 (50.84%)		263 (48.98%)		263 (48.98%)	
AE > 0.10	123 (22.91%)		117 (21.79%)		118 (21.97%)		121 (22.53%)	
Range of fitted	0.310–0.996		0.459–1.012		0.468–1.005		0.462–1.022	

Adj R^2, adjusted R^2; AE, absolute error; AIC, Akaike information criteria; BIC, Bayesian information criteria; CCS, colorectal cancer subscale; CI, confidence interval; EWB, emotional well-being; fitted, fitted values; FACT-C, Functional Assessment of Cancer Therapy-Colorectal; FWB, functional well-being; MAE, mean absolute error; Pred R^2, predicted R^2; PWB, physical well-being; RMSE, root mean square error.

* Significant with $P < 0.05$.

† Stage IV is the reference category of colorectal neoplasm staging.

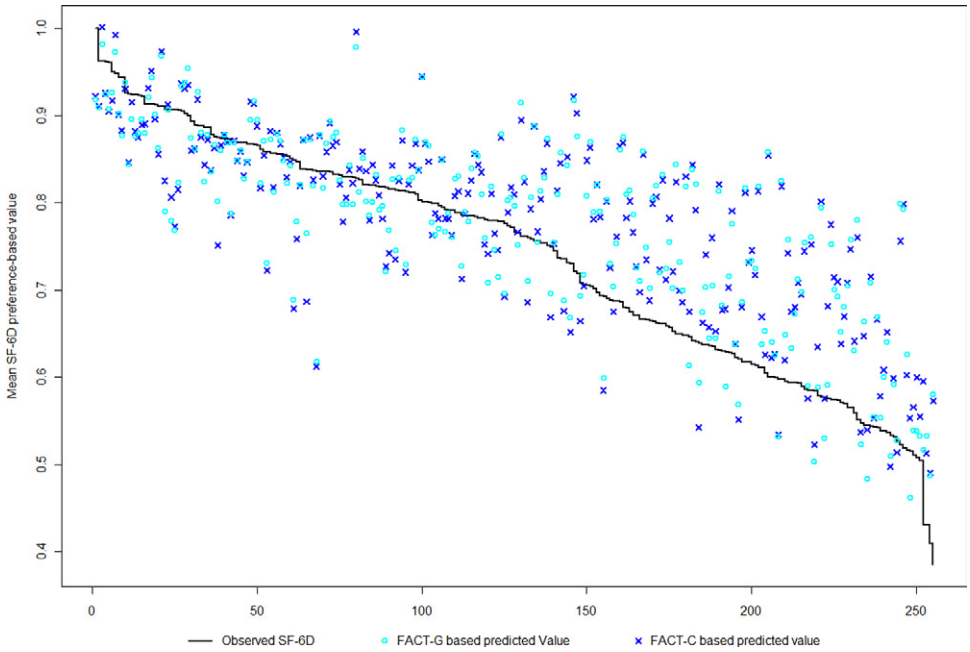

Fig. 1 – The scatter plot of mean observed and predicted SF-6D preference-based values by the health state ranking in descending order of the mean observed values.

were 0.084 and 0.065, respectively, which were better than for the FACT-G–based model. The percentage of AE greater than 0.05 was 48.98% while that of AE greater than 0.10 was 22.53%. The fitted values of SF-6D ranged from 0.462 to 1.022, the lower bound of which was 0.077 more than that of the observed range (0.385–1). The scatter plot of observed and predicted SF-6D preference-based values by FACT-C–based mapping function is also shown in Figure 1. A similar pattern of overprediction is observed for SF-6D values lower than 0.8, as demonstrated in Table 5. The mapping function of SF-6D using four scores of the FACT-C subscale (model C3) produced acceptable model performance with less than 50% of subjects with AE greater than 0.05 and the smallest Bayesian information criteria.

Model validation

Table 6 shows the prediction errors in the 10-fold cross-validation procedure. The RMSE and MAE were estimated to be 0.089 to 0.092 and 0.069 to 0.071, respectively, in FACT-G–based models and 0.088 to 0.091 and 0.068 to 0.070, respectively, in FACT-C–based models. The percentage of AE greater than 0.05 was between 50.28% and 56.04% while that of AE greater than 0.10 was between 23.07% and 24.02%. Mean SF-6D values predicted by FACT-G/FACT-C based on cross-validation were 0.825, which were consistent with the actual SF-6D values. Validation results of applying our mapping functions to patients with CRN suggested that both models predicted SF-6D preference-based values accurately.

Discussion

Mapping functions have been previously developed by using data from patients with general cancer [31, 32], esophageal cancer [33], gastric cancer [15], and prostate cancer [20]. This study is the first attempt, to our knowledge, to develop a mapping function to estimate SF-6D preference-based values from condition-specific measures for patients with CRN. The results from a relatively large sample of 537 patients with different disease severity confirmed that SF-6D preference-based values can be estimated from FACT-G/FACT-C subscale scores by using mapping functions. A 10-fold cross-validation was performed to strengthen their validity and robustness. Our findings suggested that the mapping of a more specific measure (FACT-C) had better performance than that of the general cancer measure (FACT-G), probably because it has the additional scale that captures the HRQOL domain specific to our CRN patient sample. Final models adding significant square terms, interaction terms of the subscales, and tumor stage and demographic variables into the main-effect models had slight improvements in the model performance (about 3%–6% increment in R^2, Adj R^2, and Pred R^2) at the expense of model complexity. Simpler models are sufficient to predict SF-6D preference-based values with little loss in model performance if demographic and tumor stage data are not available.

Predictive performances of the preferred models of FACT-G/FACT-C were good using statistical criteria and standards achieved by previous mapping studies. A review of mapping studies [17] found that only 4 of 30 models that reviewed mapping a condition-specific measure to a generic health preference-based measure achieved an R^2 of 60%. The R^2 value of our final FACT-G–based model was approximately 60%, and the R^2 of two of our FACT-C–based models exceeded 60%. On the whole, the R^2, Adj R^2, and Pred R^2 of all models were greater than 50%, which represented acceptable explanatory power. The predictive performance of the models was also supported by the small proportion ($<$0.10) of predictions with MAE and RMSE, and for each model less than a quarter of the sample had an AE of more than 0.10. Table 5 shows the mean error, RMSE, and MAE of predicted compared with actual preference-based values by SF-6D range for FACT-G/FACT-C–based prediction models. When comparing the ranges of the fitted values of both mapping functions to the observed range of SF-6D values, a tendency of overprediction was observed in each model for observed value of SF-6D less than 0.8 (Table 5). The size of mean error in this article varied across the SF-6D range, with larger errors when the SF-6D value was below 0.8. The mean errors in that range were larger than 0.03, which was the proposed minimally important difference for SF-6D scores [34]. Overprediction of low preference-based values also occurred in the literature mapping

Table 5 – Mean error, RMSE, and MAE of predicted compared with actual SF-6D preference-based values by SF-6D range for FACT-G– and FACT-C–based prediction models.

	FACT-G–based model			FACT-C–based model			
	G1	G2	G3	C1	C2	C3	C4
Mean error							
0.385–0.500 (n = 4)	−0.110	−0.132	−0.129	−0.113	−0.115	−0.119	−0.115
0.501–0.600 (n = 51)	−0.090	−0.093	−0.093	−0.089	−0.092	−0.086	−0.087
0.601–0.700 (n = 61)	−0.105	−0.089	−0.083	−0.104	−0.091	−0.089	−0.083
0.701–0.800 (n = 65)	−0.049	−0.037	−0.034	−0.046	−0.038	−0.037	−0.034
0.801–0.900 (n = 137)	0.010	0.015	0.013	0.012	0.014	0.014	0.013
0.901–1.000 (n = 219)	0.060	0.051	0.049	0.058	0.051	0.049	0.048
Whole range (n = 537)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RMSE							
0.385–0.500 (n = 4)	0.144	0.145	0.137	0.137	0.135	0.127	0.125
0.501–0.600 (n = 51)	0.143	0.125	0.125	0.143	0.129	0.127	0.128
0.601–0.700 (n = 61)	0.131	0.116	0.111	0.132	0.120	0.118	0.112
0.701–0.800 (n = 65)	0.085	0.088	0.086	0.084	0.087	0.086	0.084
0.801–0.900 (n = 137)	0.053	0.064	0.063	0.053	0.061	0.061	0.061
0.901–1.000 (n = 219)	0.081	0.078	0.077	0.078	0.076	0.073	0.073
Whole range (n = 537)	0.091	0.087	0.086	0.090	0.087	0.085	0.084
MAE							
0.385–0.500 (n = 4)	0.110	0.132	0.129	0.113	0.115	0.119	0.115
0.501–0.600 (n = 51)	0.119	0.100	0.102	0.118	0.104	0.098	0.101
0.601–0.700 (n = 61)	0.119	0.100	0.095	0.119	0.104	0.102	0.097
0.701–0.800 (n = 65)	0.070	0.073	0.070	0.069	0.072	0.071	0.068
0.801–0.900 (n = 137)	0.037	0.046	0.046	0.038	0.044	0.045	0.045
0.901–1.000 (n = 219)	0.065	0.062	0.061	0.063	0.059	0.058	0.057
Whole range (n = 537)	0.070	0.067	0.067	0.069	0.067	0.065	0.065

FACT-C, Functional Assessment of Cancer Therapy-Colorectal; FACT-G, Functional Assessment of Cancer Therapy-General; MAE, mean absolute error; RMSE, root mean square error; SF-6D, six-dimensional health state short form (derived from short form 36 health survey).

onto a variety of preference-based measures including the EQ-5D questionnaire [21,35,36].

A recent literature review [17] found that MAE reached 0.19, representing around 15% of the possible range of the preference-based measure. This problem has been commonly found in the use of mapping function to estimate preference-based values rather than measuring a preference-based measure directly in the study of interest. It is important that the uncertainty around the mapped estimates should be taken into account in the cost-effectiveness model and research is ongoing in this area.

Several previous studies developed the mapping functions from a condition-specific measure onto the EQ-5D questionnaire preference-based measure by censored least absolute deviations (CLAD), tobit, and standard two-part regressions [32,37–40] because the EQ-5D questionnaire suffers from high ceiling effects. Low ceiling effect observed from our SF-6D data, however, did not support the use of these models. A study by Cheung et al. [32] that mapped FACT-G onto preference-based values measured by the EQ-5D questionnaire achieved a model with R^2 of 45.0%, which was lower than the R^2 in our FACT-G–based mapping functions. In accordance with Cheung et al. [32], the PWB, EWB, and FWB but not SWB subscale scores were the significant variables in our regression models of FACT-G–based mapping functions. Our mapping functions had better predictive performance than did the models estimated by Cheung et al. [32], which could be due to a more homogenous patient population and within-sample testing in our study. Another reason for a higher explanation performance of our models was the inclusion of clinical and demographic variables that were not included in their models. Furthermore, the ceiling effects of SF-6D are lower than those of the EQ-5D Questionnaire, which could also lead to a difference in the performance of the mapping functions [41].

Although the ordinary least square regression is the most common mapping approach [17], it is limited to the normality and homoscedasticity assumptions for residuals. Because of the ceiling effect of the SF-6D preference-based measure, a small extent

Table 6 – RMSE and MAE for FACT-G– and FACT-C–based prediction models following 10-fold cross-validation procedure.

	FACT-G–based model			FACT-C–based model			
	G1	G2	G3	C1	C2	C3	C4
RMSE	0.092	0.089	0.089	0.091	0.089	0.088	0.088
MAE	0.071	0.069	0.069	0.070	0.068	0.068	0.068
AE > 0.05	301 (56.04%)	290 (54.00%)	286 (53.26%)	300 (55.87%)	274 (51.00%)	270 (50.28%)	277 (51.58%)
AE > 0.10	129 (24.02%)	124 (23.08%)	127 (23.64%)	126 (23.45%)	124 (23.08%)	124 (23.07%)	128 (23.82%)

AE, absolute error; FACT-C, Functional Assessment of Cancer Therapy-Colorectal; FACT-G, Functional Assessment of Cancer Therapy-General; RMSE, root mean square error; MAE, mean absolute error.

of violation of the normality and homoscedasticity assumptions was indicated by the residual plots of the regression models, which might have caused some biases in our coefficient estimates. Alternative mapping methods such as response mapping [42] and probabilistic mapping [43] have been used to map the Short Form 12-Item (SF-12) onto EQ-5D Questionnaire responses by others. Each EQ-5D Questionnaire dimension is classified into three response levels to allow sufficient amount of subjects for model development. However, our SF-6D data fitted poorly to these models because only a few subjects reported the worst response level in the SF-6D dimensions: physical functioning (3, 0.5%), pain (4, 0.7%), mental health (1, 0.2%), and vitality (8, 1.5%).

Limitations

Our results from Chinese patients with CRN may not be applicable to non-Chinese populations, not least because the SF-6D values predicted by the models are based on Hong Kong preferences weights. Furthermore, our models were developed for a homogeneous CRN patient population, which may not be generalizable to other cancer patient groups. It should be with caution that the model developed in this study should be validated on another sample before fitting on other data sets. In addition, the exclusion of ill patients might be the reason for the high ceiling effect observed on the PWB subscale, leading to high mean SF-6D values of 0.825. This might also be explained that most of the patients were long-term survivors with diagnosis of neoplasms for more than 6 months.

Conclusions

Mapping of FACT-G/FACT-C subscale scores onto SF-6D preference-based values reached an acceptable degree of precision for Chinese patients with CRN. The best models for the estimation of SF-6D preference-based value included the FACT-G/FACT-C subscale scores, sex, and tumor stage. These mapping functions can be applied to FACT-G/FACT-C data sets to predict SF-6D utility values for use in cost-effectiveness analysis of medical interventions for patients with CRN. Further studies on non-Chinese populations and different cancer patient groups are needed to validate the applicability of our mapping functions before wider applications.

Acknowledgments

We express our gratitude to Joyce Sing, Mandy Tai, Winnie Chan, and Deki Pun for their data collection efforts and to Eric Wan for statistical assistance.

Source of financial support: Funding for this study was provided by Small Project Funding (Project code 200907176135) from CRCG of the University of Hong Kong and Health and Health Service Research Fund (HHSRF #08090851) of Food and Health Bureau, HKSAR.

REFERENCES

[1] Gellad ZF, Provenzale D. Colorectal cancer: national and international perspective on the burden of disease and public health impact. Gastroenterology 2010;138:2177–90.
[2] National Institute for Clinical Excellence. Guide to the Methods of Technology Appraisal (reference N1618). London: National Institute for Clinical Excellence, 2008.
[3] Brazier J, Deverill M, Green C, et al. A review of the use of health status measures in economic evaluation. Health Technol Assess 1999;3:23–56.
[4] Shiroiwa T, Fukuda T, Tsutani K. Health utility scores of colorectal cancer based on societal preference in Japan. Qual Life Res 2009;18: 1095–103.
[5] Best JH, Garrison LP, Hollingworth W, et al. Preference values associated with stage III colon cancer and adjuvant chemotherapy. Qual Life Res 2010;19:391–400.
[6] Dobrez D, Cella D, Pickard AS, et al. Estimation of patient preference-based utility weights from the Functional Assessment of Cancer Therapy-General. Value Health 2007;10:266–72.
[7] Dolan P. Modeling valuations for EuroQol health states. Med Care 1997; 35:1095–108.
[8] Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002;21:271–92.
[9] Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care 2004;42:851–9.
[10] Brazier JE, Fukuhara S, Roberts J, et al. Estimating a preference-based index from the Japanese SF-36. J Clin Epidemiol 2009;62:1323–31.
[11] Ferreira LN, Ferreira PL, Pereira LN, et al. A Portuguese value set for the SF-6D. Value Health 2010;13:624–30.
[12] Lam CLK, Brazier J, McGhee SM. Valuation of the SF-6D health states is feasible, acceptable, reliable, and valid in a Chinese population. Value Health 2008;11:295–303.
[13] Cruz LN, Camey SA, Hoffmann JF, et al. Estimating the SF-6D value set for a population-based sample of Brazilians. Value Health 2011;14(Suppl.):S108–14.
[14] Buxton MJ, Lacey LA, Feagan BG, et al. Mapping from disease-specific measures to utility: an analysis of the relationships between the Inflammatory Bowel Disease Questionnaire and Crohn's Disease Activity Index in Crohn's disease and measures of utility. Value Health 2007;10:214–20.
[15] Kontodimopoulos N, Aletras VH, Paliouras D, et al. Mapping the cancer-specific EORTC QLQ-C30 to the preference-based EQ-5D, SF-6D, and 15D instruments. Value Health 2009;12:1151–7.
[16] Sauerland S, Weiner S, Dolezalova K, et al. Mapping utility scores from a disease-specific quality-of-life measure in bariatric surgery patients. Value Health 2009;12:364–70.
[17] Brazier JE, Yang Y, Tsuchiya A, et al. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. Eur J Health Econ 2010;11:215–25.
[18] Wong CKH, Lam CLK, Law WL, et al. Validity and reliability study on traditional Chinese FACT-C in Chinese patients with colorectal neoplasm. J Eval Clin Pract. In press. DOI: 10.1111/j.1365-2753.2011.01753.x.
[19] Franks P, Lubetkin EI, Gold MR, et al. Mapping the SF-12 to the EuroQol EQ-5D index in a national US sample. Med Decis Making 2004;24:247–54.
[20] Wu EQ, Mulani P, Max HF, et al. Mapping FACT-P and EORTC QLQ-C30 to patient health status measured by EQ-5D in metastatic hormone-refractory prostate cancer patients. Value Health 2007;10:408–14.
[21] Ara R, Brazier J. Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). Value Health 2008;11:1131–43.
[22] Atkin WS, Saunders BP. Surveillance guidelines after removal of colorectal adenomatous polyps. Gut 2002;51:v6–9.
[23] Greene FL, Page DL, Fleming ID, et al. AJCC Cancer Staging Manual 6th ed.). New York: Springer, 2002.
[24] Wong CKH, Lam CLK, Poon JTC, et al. Direct medical costs of care for Chinese patients with colorectal neoplasia: a health care service provider perspective. J Eval Clin Pract. In press. DOI: 10.1111/j.1365-2753.2011.01776.x.
[25] Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) measurement system: properties, applications, and interpretation. Health Qual Life Outcomes 2003;1:79.
[26] Ward WL, Hahn EA, Mo F, et al. Reliability and validity of the Functional Assessment of Cancer Therapy-Colorectal (FACT-C) quality of life instrument. Qual Life Res 1999;8:181–95.
[27] Hahn EA, Rao D, Cella D, et al. Comparability of interview- and self-administration of the Functional Assessment of Cancer Therapy-General (FACT-G) in English- and Spanish-speaking ambulatory cancer patients. Med Care 2008;46:423–31.
[28] McGhee SM, Brazier J, Lam CLK, et al. Quality adjusted life years: population-specific measurement of the quality component. Hong Kong Med J 2011;17(Suppl 6):S17–21.
[29] Lam ETP, Lam CCK, Lai CL, et al. Health-related quality of life of Southern Chinese with chronic hepatitis B infection. Health Qual Life Outcomes 2009;7:52.
[30] Montgomery DC, Peck EA, Vining GG. Diagnostics for Leverage and Influence: Introduction to Linear Regression Analysis (4th ed.). Hoboken, NJ: Wiley-Interscience, 2006.
[31] Pickard AS, Shaw JW, Lin HW, et al. A patient-based utility measure of health for clinical trials of cancer therapy based on the European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire. Value Health 2009;12(Suppl.):977–88.
[32] Cheung YB, Thumboo J, Gao F, et al. Mapping the English and Chinese versions of the Functional Assessment of Cancer Therapy-General to the EQ-5D utility index. Value Health 2009;12:371–6.
[33] McKenzie L, van der Pol M. Mapping the EORTC QLQ C-30 onto the EQ-5D instrument: the potential to estimate QALYs without generic preference data. Value Health 2009;12:167–71.

[34] Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res 2005;14:1523–32.
[35] Versteegh M, Rowen DL, Brazier JE, et al. Mapping onto EQ-5D for patients in poor health. Health Qual Life Outcomes 2010;8:141.
[36] Rowen DL, Brazier JE, Roberts J. Mapping SF-36 onto the EQ-5D index: how reliable is the relationship? Health Qual Life Outcomes 2009;7:27.
[37] Payakachat N, Summers K, Pleil A, et al. Predicting EQ-5D utility scores from the 25-item National Eye Institute Vision Function Questionnaire (NEI-VFQ 25) in patients with age-related macular degeneration. Qual Life Res 2009;18:801–13.
[38] Xie F, Pullenayegum EM, Li SC, et al. Use of a disease-specific instrument in economic evaluations: mapping WOMAC onto the EQ-5D utility index. Value Health 2010;13:873–8.
[39] Sullivan PW, Ghushchyan V. Mapping the EQ-5D index from the SF-12: US general population preferences in a nationally representative sample. Med Decis Making 2006;26:401–9.
[40] Huang IC, Frangakis C, Atkinson MJ, et al. Addressing ceiling effects in health status measures: a comparison of techniques applied to measures for people with HIV disease. Health Serv Res 2008;43:327–39.
[41] Brazier J, Roberts J, Tsuchiya A, et al. A comparison of the EQ-5D and SF-6D across seven patient groups. Health Econ 2004;13:873–84.
[42] Gray AM, Rivero-Arias O, Clarke PM. Estimating the association between SF-12 responses and EQ-5D utility values by response mapping. Med Decis Making 2006;26:18–29.
[43] Le QA, Doctor JN. Probabilistic mapping of descriptive health status responses onto health state utilities using Bayesian networks: an empirical analysis converting SF-12 into EQ-5D utility index in a national US sample. Med Care 2011;49:451–60.